



DNA Sequences Encoding Novel Growth/
Differentiation Factors

The present invention relates to DNA sequences encoding novel growth/differentiation factors of the TGF- β family. In particular, it relates to novel DNA sequences encoding TGF- β -like proteins, to the isolation of said DNA sequences, to expression plasmids containing said DNA, to microorganisms transformed by said expression plasmid, to the production of said protein by culturing said transformant, and to pharmaceutical compositions containing said protein. The TGF- β family of growth factors comprising BMP, TGF, and Inhibin related proteins (Roberts and Sporn, Handbook of Experimental Pharmacology 95 (1990), 419-472) is of particular relevance in a wide range of medical treatments and applications. These factors are useful in processes relating to wound healing and tissue repair. Furthermore, several members of the TGF- β family are tissue inductive, especially osteo-inductive, and consequently play a crucial role in inducing cartilage and bone development.

Wozney, Progress in Growth Factor Research 1 (1989), 267-280 and Vale et al., Handbook of Experimental Pharmacology 95 (1990), 211-248 describe different growth factors such as those relating to the BMP (bone morphogenetic proteins) and the Inhibin group. The members of these groups share significant structural similarity. The precursor of the protein is composed of an aminoterminal signal sequence, a propeptide and a carboxyterminal sequence of about 110 amino acids, which is subsequently cleaved from the precursor and represents the mature protein. Furthermore, their members are defined by virtue of amino acid sequence homology. The

mature protein contains the most conserved sequences, especially seven cysteine residues which are conserved among the family members. The TGF- β -like proteins are multifunctional, hormonally active growth factors. They also share related biological activities such as chemotactic attraction of cells, promoting cell differentiation and their tissue-inducing capacity, such as cartilage- and bone-inducing capacity. U.S. Patent No. 5,013,649 discloses DNA sequences encoding osteo-inductive proteins termed BMP-2 proteins (bone morphogenetic protein), and U.S. patent applications serial nos. 179 101 and 179 197 disclose the BMP proteins BMP-1 and BMP-3. Furthermore, many cell types are able to synthesize TGF- β -like proteins and virtually all cells possess TGF- β receptors.

RECEIVED
COURT EXHIBIT
DEPT. OF JUSTICE

Taken together, these proteins show differences in their structure, leading to considerable variation in their detailed biological function. Furthermore, they are found in a wide variety of different tissues and developmental stages. Consequently, they might possess differences concerning their function in detail, for instance the required cellular physiological environment, their lifespan, their targets, their requirement for accessory factors, and their resistance to degradation. Thus, although numerous proteins exhibiting tissue-inductive, especially osteo-inductive potential are described, their natural role in the organism and, more importantly, their medical relevance must still be elucidated in detail. The occurrence of still-unknown members of the TGF- β family relevant for osteogenesis or differentiation/induction of other tissues is strongly suspected. However, a major problem in the isolation of these new TGF- β -like proteins is that their functions cannot yet be described precisely enough for the design of a discriminative bioassay. On the other hand, the expected nucleotide sequence homology to known members of the family would be too low to

allow for screening by classical nucleic acid hybridization techniques. Nevertheless, the further isolation and characterization of new TGF- β -like proteins is urgently needed in order to get hold of the whole set of induction and differentiation proteins meeting all desired medical requirements. These factors might find useful medical applications in defect healing and treatments of degenerative disorders of bone and/or other tissues like, for example, kidney and liver.

Thus, the technical problem underlying the present invention essentially is to provide DNA sequences coding for new members of the TGF- β protein family having mitogenic and/or differentiation-inductive, e.g. osteo-inductive potential.

The solution to the above technical problem is achieved by providing the embodiments characterized in claims 1 to 17. Other features and advantages of the invention will be apparent from the description of the preferred embodiments and the drawings. The sequence listings and drawings will now briefly be described.

SEQ ID NO. 1 shows the nucleotide sequence of MP-52, i.e. the embryo derived sequence corresponding to the mature peptide and most of the sequence coding for the propeptide of MP-52.

Some of the propeptide sequence at the 5'-end of MP-52 has not been characterized so far.

SEQ ID NO. 2 shows the nucleotide sequence of MP-121, i.e. the liver derived sequence corresponding to the mature peptide, the sequence coding for the propeptide of MP-121, and sequences 5' and 3' to the coding region.

The start codon begins with nucleotide 128 of SEQ ID NO.2. The sequence coding for the mature MP121 polypeptide begins with nucleotide 836 of SEQ ID NO. 2. The stop codon begins with nucleotide 1184 of SEQ ID NO. 2. The sequence coding for the precursor protein has a length of 1056 bp. The sequence coding for the propeptide has a length of 708 bp and the sequence coding for the mature peptide has a length of 348 bp.

SEQ ID NO. 3 shows the amino acid sequence of MP-52 as deduced from SEQ ID NO. 1.

SEQ ID NO. 4 shows the amino acid sequence of MP-121 as deduced from sequence SEQ ID NO.2. The sequence of the mature polypeptide begins with amino acid 237 of SEQ ID NO. 4. The precursor protein has a length of 352 amino acids. The propeptide and the mature peptide have a length of 236 and 116 amino acids, respectively.

SEQ ID NO. 5 shows a part of the nucleotide sequence of the liver derived sequence of MP-121.

SEQ ID NO. 6 shows a part of the nucleotide sequence of the embryo derived sequence of MP-52.

The shorter DNA-sequences SEQ ID NO. 5 and 6 can be useful for example for isolation of further members of the TGF- β -protein family.

~~Figure 1 shows an alignment of the amino acid sequences of MP-52 and MP-121 starting from the first of the seven conserved cysteines with some related proteins. 1a shows the alignment of MP-52 with some members of the BMP protein family; 1b shows the alignment of MP-121 with some members of the Inhibin protein family. * indicates that the amino acid~~

b17 is the same in all proteins compared; + indicates that the amino acid is the same in at least one of the proteins compared with MP-52 (Fig. 1a) or MP-121 (Fig. 1b).

~~Sequence 2~~ Figure 2 shows the nucleotide sequences of the oligonucleotide primer as used in the present invention and an alignment of these sequences with known members of the TGF- β family. M means A or C; S means C or G; R means A or G; and K means G or T. 2a depicts the sequence of the primer OD; 2b shows the sequence of the primer OID.

The present invention relates to novel TGF- β -like proteins and provides DNA sequences contained in the corresponding genes. Such sequences include nucleotide sequences comprising the sequence

ATGAACTCCATGGACCCCGAGTCCACA and
CTTCTCAAGGCCAACACAGCTGCAGGCACC

and in particular sequences as illustrated in SEQ ID Nos. 1 and 2, allelic derivatives of said sequences and DNA sequences degenerated as a result of the genetic code for said sequences. They also include DNA sequences hybridizing under stringent conditions with the DNA sequences mentioned above and containing the following amino acid sequences:

Met-Asn-Ser-Met-Asp-Pro-Glu-Ser-Thr or
Leu-Leu-Lys-Ala-Asn-Thr-Ala-Ala-Gly-Thr.

Although said allelic, degenerate and hybridizing sequences may have structural divergencies due to naturally occurring mutations, such as small deletions or substitutions, they will usually still exhibit essentially the same useful properties, allowing their use in basically the same medical applications.

According to the present invention, the term "hybridization" means conventional hybridization conditions, preferably

conditions with a salt concentration of 6 x SSC at 62° to 66°C followed by a one-hour wash with 0.6 x SSC, 0.1% SDS at 62° to 66°C. The term "hybridization" preferably refers to stringent hybridization conditions with a salt concentration of 4 x SSC at 62°-66°C followed by a one-hour wash with 0.1 x SSC, 0.1% SDS at 62°-66°C.

Important biological activities of the encoded proteins, preferably MP-52, comprise a mitogenic and osteo-inductive potential and can be determined in assays according to Seyedin et al., PNAS 82 (1985), 2267-2271 or Sampath and Reddi, PNAS 78 (1981), 7599-7603.

The biological properties of the proteins according to the invention, preferably MP-121, may be determined, e.g., by means of the assays according to Wrana et al. (Cell 71, 1003-1014 (1992)), Ling et al. (Proc. Natl. Acad. of Science, 82, 7217-7221 (1985)), Takuwa et al. (Am. J. Physiol., 257, E797-E803 (1989)), Fann and Patterson (Proc. Natl. Acad. of Science, 91, 43-47 (1994)), Broxmeyer et al. (Proc. Natl. Acad. of Science, 85, 9052-9056 (1988)), Green et al. (Cell, 71, 731-739 (1992)), Partridge et al. (Endocrinology, 108, 213-219 (1981)) or Roberts et al. (PNAS 78, 5339-5343 (1981)).

Preferred embodiments of the present invention are DNA sequences as defined above and obtainable from vertebrates, preferably mammals such as pig or cow and from rodents such as rat or mouse, and in particular from primates such as humans.

Particularly preferred embodiments of the present invention are the DNA sequences termed MP-52 and MP-121 which are shown in SEQ ID Nos. 1 and 2. The corresponding transcripts of MP-52 were obtained from embryogenic tissue and code for a

protein showing considerable amino acid homology to the mature part of the BMP-like proteins (see Fig. 1a). The protein sequences of BMP2 (=BMP2A) and BMP4 (=BMP2B) are described in Wozney et al., Science Vol 242, 1528-1534 (1988). The respective sequences of BMP5, BMP6 and BMP7 are described in Celeste et al., Proc.Natl.Acad.Sci. USA Vol 87, 9843-9847 (1990). Some typical sequence homologies, which are specific to known BMP-sequences only, were also found in the propeptide part of MP-52, whereas other parts of the precursor part of MP-52 show marked differences to BMP-precursors. The mRNA of MP-121 was detected in liver tissue, and its corresponding amino acid sequence shows homology to the amino acid sequences of the Inhibin protein chains (see Fig. 1b). cDNA sequences encoding TGF- β -like proteins have not yet been isolated from liver tissue, probably due to a low abundance of TGF- β specific transcripts in this tissue. In embryogenic tissue, however, sequences encoding known TGF- β -like proteins can be found in relative abundance. The inventors have recently detected the presence of a collection of TGF- β -like proteins in liver as well. The high background level of clones related to known factors of this group presents the main difficulty in establishing novel TGF- β -related sequences from these and probably other tissues. In the present invention, the cloning was carried out according to the method described below. Once the DNA sequence has been cloned, the preparation of host cells capable of producing the TGF- β -like proteins and the production of said proteins can be easily accomplished using known recombinant DNA techniques comprising constructing the expression plasmids encoding said protein and transforming a host cell with said expression plasmid, cultivating the transformant in a suitable culture medium, and recovering the product having TGF- β -like activity.

Thus, the invention also relates to recombinant molecules comprising DNA sequences as described above, optionally linked to an expression control sequence. Such vectors may be useful in the production of TGF- β -like proteins in stably or transiently transformed cells. Several animal, plant, fungal and bacterial systems may be employed for the transformation and subsequent cultivation process. Preferably, expression vectors which can be used in the invention contain sequences necessary for the replication in the host cell and are autonomously replicable. It is also preferable to use vectors containing selectable marker genes which can be easily selected for transformed cells. The necessary operation is well-known to those skilled in the art.

It is another object of the invention to provide a host cell transformed by an expression plasmid of the invention and capable of producing a protein of the TGF- β family. Examples of suitable host cells include various eukaryotic and prokaryotic cells, such as *E. coli*, insect cells, plant cells, mammalian cells, and fungi such as yeast.

Another object of the present invention is to provide a protein of the TGF- β family encoded by the DNA sequences described above and displaying biological features such as tissue-inductive, in particular osteo-inductive and/or mitogenic capacities possibly relevant to therapeutical treatments. The above-mentioned features of the protein might vary depending upon the formation of homodimers or heterodimers. Such structures may prove useful in clinical applications as well. The amino acid sequence of the especially preferred proteins of the TGF- β -family (MP-52 and MP-121) are shown in SEQ ID NO. 3 and SEQ ID NO. 4.

It is a further aspect of the invention to provide a process for the production of TGF- β -like proteins. Such a process

055056 055057 055058
055059 055060 055061
055062 055063 055064
055065 055066 055067
055068 055069 055070
055071 055072 055073
055074 055075 055076
055077 055078 055079
055080 055081 055082
055083 055084 055085
055086 055087 055088
055089 055090 055091
055092 055093 055094
055095 055096 055097
055098 055099 055100
055101 055102 055103
055104 055105 055106
055107 055108 055109
055110 055111 055112
055113 055114 055115
055116 055117 055118
055119 055120 055121
055122 055123 055124
055125 055126 055127
055128 055129 055130
055131 055132 055133
055134 055135 055136
055137 055138 055139
055140 055141 055142
055143 055144 055145
055146 055147 055148
055149 055150 055151
055152 055153 055154
055155 055156 055157
055158 055159 055160
055161 055162 055163
055164 055165 055166
055167 055168 055169
055170 055171 055172
055173 055174 055175
055176 055177 055178
055179 055180 055181
055182 055183 055184
055185 055186 055187
055188 055189 055190
055191 055192 055193
055194 055195 055196
055197 055198 055199
055200 055201 055202
055203 055204 055205
055206 055207 055208
055209 055210 055211
055212 055213 055214
055215 055216 055217
055218 055219 055220
055221 055222 055223
055224 055225 055226
055227 055228 055229
055230 055231 055232
055233 055234 055235
055236 055237 055238
055239 055240 055241
055242 055243 055244
055245 055246 055247
055248 055249 055250
055251 055252 055253
055254 055255 055256
055257 055258 055259
055260 055261 055262
055263 055264 055265
055266 055267 055268
055269 055270 055271
055272 055273 055274
055275 055276 055277
055278 055279 055280
055281 055282 055283
055284 055285 055286
055287 055288 055289
055290 055291 055292
055293 055294 055295
055296 055297 055298
055299 055300 055301
055302 055303 055304
055305 055306 055307
055308 055309 055310
055311 055312 055313
055314 055315 055316
055317 055318 055319
055320 055321 055322
055323 055324 055325
055326 055327 055328
055329 055330 055331
055332 055333 055334
055335 055336 055337
055338 055339 055340
055341 055342 055343
055344 055345 055346
055347 055348 055349
055350 055351 055352
055353 055354 055355
055356 055357 055358
055359 055360 055361
055362 055363 055364
055365 055366 055367
055368 055369 055370
055371 055372 055373
055374 055375 055376
055377 055378 055379
055380 055381 055382
055383 055384 055385
055386 055387 055388
055389 055390 055391
055392 055393 055394
055395 055396 055397
055398 055399 055400
055401 055402 055403
055404 055405 055406
055407 055408 055409
055410 055411 055412
055413 055414 055415
055416 055417 055418
055419 055420 055421
055422 055423 055424
055425 055426 055427
055428 055429 055430
055431 055432 055433
055434 055435 055436
055437 055438 055439
055440 055441 055442
055443 055444 055445
055446 055447 055448
055449 055450 055451
055452 055453 055454
055455 055456 055457
055458 055459 055460
055461 055462 055463
055464 055465 055466
055467 055468 055469
055470 055471 055472
055473 055474 055475
055476 055477 055478
055479 055480 055481
055482 055483 055484
055485 055486 055487
055488 055489 055490
055491 055492 055493
055494 055495 055496
055497 055498 055499
055400 055401 055402
055403 055404 055405
055406 055407 055408
055409 055410 055411
055412 055413 055414
055415 055416 055417
055418 055419 055420
055421 055422 055423
055424 055425 055426
055427 055428 055429
055430 055431 055432
055433 055434 055435
055436 055437 055438
055439 055440 055441
055442 055443 055444
055445 055446 055447
055448 055449 055450
055451 055452 055453
055454 055455 055456
055457 055458 055459
055460 055461 055462
055463 055464 055465
055466 055467 055468
055469 055470 055471
055472 055473 055474
055475 055476 055477
055478 055479 055480
055481 055482 055483
055484 055485 055486
055487 055488 055489
055490 055491 055492
055493 055494 055495
055496 055497 055498
055499 055500 055501
055502 055503 055504
055505 055506 055507
055508 055509 055510
055511 055512 055513
055514 055515 055516
055517 055518 055519
055520 055521 055522
055523 055524 055525
055526 055527 055528
055529 055530 055531
055532 055533 055534
055535 055536 055537
055538 055539 055540
055541 055542 055543
055544 055545 055546
055547 055548 055549
055550 055551 055552
055553 055554 055555
055556 055557 055558
055559 055560 055561
055562 055563 055564
055565 055566 055567
055568 055569 055570
055571 055572 055573
055574 055575 055576
055577 055578 055579
055580 055581 055582
055583 055584 055585
055586 055587 055588
055589 055590 055591
055592 055593 055594
055595 055596 055597
055598 055599 055600
055601 055602 055603
055604 055605 055606
055607 055608 055609
055610 055611 055612
055613 055614 055615
055616 055617 055618
055619 055620 055621
055622 055623 055624
055625 055626 055627
055628 055629 055630
055631 055632 055633
055634 055635 055636
055637 055638 055639
055640 055641 055642
055643 055644 055645
055646 055647 055648
055649 055650 055651
055652 055653 055654
055655 055656 055657
055658 055659 055660
055661 055662 055663
055664 055665 055666
055667 055668 055669
055670 055671 055672
055673 055674 055675
055676 055677 055678
055679 055680 055681
055682 055683 055684
055685 055686 055687
055688 055689 055690
055691 055692 055693
055694 055695 055696
055697 055698 055699
055600 055601 055602
055603 055604 055605
055606 055607 055608
055609 055610 055611
055612 055613 055614
055615 055616 055617
055618 055619 055620
055621 055622 055623
055624 055625 055626
055627 055628 055629
055630 055631 055632
055633 055634 055635
055636 055637 055638
055639 055640 055641
055642 055643 055644
055645 055646 055647
055648 055649 055650
055651 055652 055653
055654 055655 055656
055657 055658 055659
055660 055661 055662
055663 055664 055665
055666 055667 055668
055669 055670 055671
055672 055673 055674
055675 055676 055677
055678 055679 055680
055681 055682 055683
055684 055685 055686
055687 055688 055689
055690 055691 055692
055693 055694 055695
055696 055697 055698
055699 055700 055701
055702 055703 055704
055705 055706 055707
055708 055709 055710
055711 055712 055713
055714 055715 055716
055717 055718 055719
055720 055721 055722
055723 055724 055725
055726 055727 055728
055729 055730 055731
055732 055733 055734
055735 055736 055737
055738 055739 055740
055741 055742 055743
055744 055745 055746
055747 055748 055749
055750 055751 055752
055753 055754 055755
055756 055757 055758
055759 055760 055761
055762 055763 055764
055765 055766 055767
055768 055769 055770
055771 055772 055773
055774 055775 055776
055777 055778 055779
055780 055781 055782
055783 055784 055785
055786 055787 055788
055789 055790 055791
055792 055793 055794
055795 055796 055797
055798 055799 055800
055801 055802 055803
055804 055805 055806
055807 055808 055809
055810 055811 055812
055813 055814 055815
055816 055817 055818
055819 055820 055821
055822 055823 055824
055825 055826 055827
055828 055829 055830
055831 055832 055833
055834 055835 055836
055837 055838 055839
055840 055841 055842
055843 055844 055845
055846 055847 055848
055849 055850 055851
055852 055853 055854
055855 055856 055857
055858 055859 055860
055861 055862 055863
055864 055865 055866
055867 055868 055869
055870 055871 055872
055873 055874 055875
055876 055877 055878
055879 055880 055881
055882 055883 055884
055885 055886 055887
055888 055889 055890
055891 055892 055893
055894 055895 055896
055897 055898 055899
055800 055801 055802
055803 055804 055805
055806 055807 055808
055809 055810 055811
055812 055813 055814
055815 055816 055817
055818 055819 055820
055821 055822 055823
055824 055825 055826
055827 055828 055829
055830 055831 055832
055833 055834 055835
055836 055837 055838
055839 055840 055841
055842 055843 055844
055845 055846 055847
055848 055849 055850
055851 055852 055853
055854 055855 055856
055857 055858 055859
055860 055861 055862
055863 055864 055865
055866 055867 055868
055869 055870 055871
055872 055873 055874
055875 055876 055877
055878 055879 055880
055881 055882 055883
055884 055885 055886
055887 055888 055889
055890 055891 055892
055893 055894 055895
055896 055897 055898
055899 055900 055901
055902 055903 055904
055905 055906 055907
055908 055909 0559010
0559011 0559012 0559013
0559014 0559015 0559016
0559017 0559018 0559019
0559020 0559021 0559022
0559023 0559024 0559025
0559026 0559027 0559028
0559029 0559030 0559031
0559032 0559033 0559034
0559035 0559036 0559037
0559038 0559039 0559040
0559041 0559042 0559043
0559044 0559045 0559046
0559047 0559048 0559049
0559050 0559051 0559052
0559053 0559054 0559055
0559056 0559057 0559058
0559059 0559060 0559061
0559062 0559063 0559064
0559065 0559066 0559067
0559068 0559069 0559070
0559071 0559072 0559073
0559074 0559075 0559076
0559077 0559078 0559079
0559080 0559081 0559082
0559083 0559084 0559085
0559086 0559087 0559088
0559089 0559090 0559091
0559092 0559093 0559094
0559095 0559096 0559097
0559098 0559099 0559100
0559101 0559102 0559103
0559104 0559105 0559106
0559107 0559108 0559109
0559110 0559111 0559112
0559113 0559114 0559115
0559116 0559117 0559118
0559119 0559120 0559121
0559122 0559123 0559124
0559125 0559126 0559127
0559128 0559129 0559130
0559131 0559132 0559133
0559134 0559135 0559136
0559137 0559138 0559139
0559140 0559141 0559142
0559143 0559144 0559145
0559146 0559147 0559148
0559149 0559150 0559151
0559152 0559153 0559154
0559155 0559156 0559157
0559158 0559159 0559160
0559161 0559162 0559163
0559164 0559165 0559166
0559167 0559168 0559169
0559170 0559171 0559172
0559173 0559174 0559175
0559176 0559177 0559178
0559179 0559180 0559181
0559182 0559183 0559184
0559185 0559186 0559187
0559188 0559189 0559190
0559191 0559192 0559193
0559194 0559195 0559196
0559197 0559198 0559199
0559200 0559201 0559202
0559203 0559204 0559205
0559206 0559207 0559208
0559209 0559210 0559211
0559212 0559213 0559214
0559215 0559216 0559217
0559218 0559219 0559220
0559221 0559222 0559223
0559224 0559225 0559226
0559227 0559228 0559229
0559230 0559231 0559232
0559233 0559234 0559235
0559236 0559237 0559238
0559239 0559240 0559241
0559242 0559243 0559244
0559245 0559246 0559247
0559248 0559249 0559250
0559251 0559252 0559253
0559254 0559255 0559256
0559257 0559258 0559259
0559260 0559261 0559262
0559263 0559264 0559265
0559266 0559267 0559268
0559269 0559270 0559271
0559272 0559273 0559274
0559275 0559276 0559277
0559278 0559279 0559280
0559281 0559282 0559283
0559284 0559285 0559286
0559287 0559288 0559289
05

comprises cultivating a host cell being transformed with a DNA sequence of the present invention in a suitable culture medium and purifying the TGF- β -like protein produced. Thus, this process will allow the production of a sufficient amount of the desired protein for use in medical treatments or in applications using cell culture techniques requiring growth factors for their performance. The host cell is obtainable from bacteria such as *Bacillus* or *Escherichia coli*, from fungi such as yeast, from plants such as tobacco, potato, or *Arabidopsis*, and from animals, in particular vertebrate cell lines such as the Mo-, COS- or CHO cell line.

DECODED
DECODED
DECODED
DECODED
DECODED
DECODED
DECODED

Yet another aspect of the present invention is to provide a particularly sensitive process for the isolation of DNA sequences corresponding to low abundance mRNAs in the tissues of interest. The process of the invention comprises the combination of four different steps. First, the mRNA has to be isolated and used in an amplification reaction using oligonucleotide primers. The sequence of the oligonucleotide primers contains degenerated DNA sequences derived from the amino acid sequence of proteins related to the gene of interest. This step may lead to the amplification of already known members of the gene family of interest, and these undesired sequences would therefore have to be eliminated. This object is achieved by using restriction endonucleases which are known to digest the already-analyzed members of the gene family. After treatment of the amplified DNA population with said restriction endonucleases, the remaining desired DNA sequences are isolated by gel electrophoresis and reamplified in a third step by an amplification reaction, and in a fourth step they are cloned into suitable vectors for sequencing. To increase the sensitivity and efficiency, steps two and three are repeatedly performed, at least two times in one embodiment of this process.

In a preferred embodiment, the isolation process described above is used for the isolation of DNA sequences from liver tissue. In a particularly preferred embodiment of the above-described process, one primer used for the PCR experiment is homologous to the polyA tail of the mRNA, whereas the second primer contains a gene-specific sequence. The techniques employed in carrying out the different steps of this process (such as amplification reactions or sequencing techniques) are known to the person skilled in the art and described, for instance, in Sambrook et al., 1989, "Molecular Cloning: A laboratory manual", Cold Spring Harbor Laboratory Press.

It is another object of the present invention to provide pharmaceutical compositions containing a therapeutically-effective amount of a protein of the TGF- β family of the present invention. Optionally, such a composition comprises a composition can be used in wound healing and tissue repair as well as in the healing of bone, cartilage, or tooth defects, either individually or in conjunction with suitable carriers, and possibly with other related proteins or growth factors. Thus, a therapeutic composition of the invention may include, but is not limited to, the MP-52 encoded protein in conjunction with the MP-121 encoded protein, and optionally with other known biologically-active substances such as EGF (epidermal growth factor) or PDGF (platelet derived growth factor). Another possible clinical application of a TGF- β -like protein is the use as a suppressor of the immuno-response, which would prevent rejection of organ transplants. The pharmaceutical composition comprising the proteins of the invention can also be used prophylactically, or can be employed in cosmetic plastic surgery. Furthermore, the application of the composition is not limited to humans but can include animals, in particular domestic animals, as well. Possible applications of the pharmaceutical composition

according to the invention include furthermore treatment or prevention of connective tissue, skin, mucous membrane, endothelial, epithelial, neuronal or renal defects, use in the case of dental implants, use as a morphogenic factor used for inducing liver tissue growth, induction of the proliferation of precursor cells or bone marrow cells, for maintaining a differentiated state and the treatment of impaired fertility or for contraception.

Finally, another object of the present invention is an antibody or antibody fragment, which is capable of specifically binding to the proteins of the present invention. Methods to raise such specific antibody are general knowledge. Preferably such an antibody is a monoclonal antibody. Such antibodies or antibody fragments might be useful for diagnostic methods.

The following examples illustrate in detail the invention disclosed, but should not be construed as limiting the invention.

Example 1

Isolation of MP-121

- 1.1 Total RNA was isolated from human liver tissue (40-year-old-male) by the method of Chirgwin et al., Biochemistry 18 (1979), 5294-5299. Poly A⁺ RNA was separated from total RNA by oligo (dT) chromatography according to the instructions of the manufacturer (Stratagene Poly (A) Quick columns).
- 1.2 For the reverse transcription reaction, poly A⁺ RNA (1-2.5 µg) derived from liver tissue was heated for 5 minutes to 65°C and cooled rapidly on ice. The reverse transcription reagents containing 27 U RNA guard

6647260 85510650

(Pharmacia), 2.5 µg oligo d(T)₁₂₋₁₈ (Pharmacia) 5 x buffer (250 mM Tris/HCl pH 8.5; 50 mM MgCl₂; 50 mM DTT; 5 mM each dNTP; 600 mM KCl) and 20 units avian myeloblastosis virus reverse transcriptase (AMV, Boehringer Mannheim) per µg poly (A⁺) RNA were added. The reaction mixture (25 µl) was incubated for 2 hours at 42°C. The liver cDNA pool was stored at -20°C.

- 1.3 The deoxynucleotide primers OD and OID (Fig. 2) designed to prime the amplification reaction were generated on an automated DNA-synthesizer (Biosearch). Purification was done by denaturating polyacrylamide gel electrophoresis and isolation of the main band from the gel by isotachophoresis. The oligonucleotides were designed by aligning the nucleic acid sequences of some known members of the TGF-β family and selecting regions of the highest conservation. An alignment of this region is shown in Fig. 2. In order to facilitate cloning, both oligonucleotides contained EcoR I restriction sites and OD additionally contained an Nco I restriction site at its 5' terminus.
- 1.4 In the polymerase chain reaction, a liver-derived cDNA pool was used as a template in a 50 µl reaction mixture. The amplification was performed in 1 x PCR-buffer (16.6 mM (NH₄)₂SO₄; 67 mM Tris/HCl pH 8.8; 2 mM MgCl₂; 6.7 µM EDTA; 10 mM β-mercaptoethanol; 170 µg/ml BSA (Gibco)), 200 µM each dNTP (Pharmacia), 30 pmol each oligonucleotide (OD and OID) and 1.5 units Taq polymerase (AmpliTaq, Perkin Elmer Cetus). The PCR reaction contained cDNA corresponding to 30 ng of poly (A⁺) RNA as starting material. The reaction mixture was overlayed by paraffine and 40 cycles (cycle 1: 80s 93°C/40s 52°C/40s 72°C; cycles 2-9: 60s 93°C/40s 52°C/40s 72°C; cycles 10-29: 60s 93°C/40s 52°C/60s

6641260 "SISTEMO"

72°C; cycles 30-31: 60s 93°C/40s 52°C/90s 72°C; cycle 40: 60s 93°C/40s 52°C/420s 72°C) of the PCR were performed. Six PCR-reaction mixtures were pooled, purified by subsequent extractions with equal volumes of phenol, phenol/chloroform (1:1 (v/v)) and chloroform/isoamylalcohol (24:1 (v/v)) and concentrated by ethanol precipitation.

- 1.5 One half of the obtained PCR pool was sufficient for digestion with the restriction enzymes Sph I (Pharmacia) and AlwN I (Biolabs). The second half was digested in a series of reactions by the restriction enzymes Ava I (BRL), AlwN I (Biolabs) and Tfi I (Biolabs). The restriction endonuclease digestions were performed in 100 µl at 37°C (except Tfi I at 65°C) using 8 units of each enzyme in a 2- to 12-hour reaction in a buffer recommended by the manufacturer.
- 1.6 Each DNA sample was fractionated by electrophoresis using a 4% agarose gel (3% FMC Nusieve agarose, Biozym and 1% agarose, BRL) in Tris borate buffer (89 mM Trisbase, 89 mM boric acid, 2 mM EDTA, pH 8). After ethidiumbromide staining uncleaved amplification products (about 200 bp; size marker was run in parallel) were excised from the gel and isolated by phenol extraction: an equal volume of phenols was added to the excised agarose, which was minced to small pieces, frozen for 10 minutes, vortexed and centrifuged. The aqueous phase was collected, the interphase reextracted by the same volume TE-buffer, centrifuged and both aqueous phases were combined. DNA was further purified twice by phenol/chloroform and once by chloroform/isoamylalcohol extraction.
- 1.7 After ethanol precipitation, one fourth or one fifth of the isolated DNA was reamplified using the same

52°C/60s 72°C; cycle 13: 60s 93°C/40s 52°C/420s
72°C). The reamplification products were purified,
restricted with the same enzymes as above and the
uncleaved products were isolated from agarose gels as
mentioned above for the amplification products. The
reamplification followed by restriction and gel
isolation was repeated once.

- 1.8 After the last isolation from the gel, the amplification products were digested by 4 units EcoR I (Pharmacia) for 2 hours at 37°C using the buffer recommended by the manufacturer. One fourth of the restriction mixture was ligated to the vector pBluescriptII SK+ (Stratagene) which was digested likewise by EcoR I. After ligation, 24 clones from each enzyme combination were further analyzed by sequence analysis. The sample restricted by AlwN I and Sph I contained no new sequences, only BMP6 and Inhibin β A sequences. 19 identical new sequences, which were named MP-121, were found by the Ava I, AlwN I and Tfi I restricted samples. The MP-121 containing plasmids were called pSK MP-121 (OD/OID). One sequence differed from this mainly-found sequence by two nucleotide exchanges. Ligation reaction and transformation in *E. coli* HB101 were performed as described in Sambrook et al., Molecular cloning: A laboratory manual (1989). Transformants were selected by Ampicillin resistance and the plasmid DNAs were isolated according to standard protocols (Sambrook et al. (1989)). Analysis was done by sequencing the double-stranded plasmids by "dideoxyribonucleotide chain termination sequencing" with the sequencing kit "Sequenase Version 2.0" (United States Biochemical Corporation).

O P E N D O N G R A F I C O

S105 The clone was completed to the 3' end of the c-DNA by a method described in detail by Frohman (Amplifications, published by Perkin-Elmer Corporation, issue 5 (1990), pp 11-15). The same liver mRNA which was used for the isolation of the first fragment of MP-121 was reverse transcribed using a primer consisting of oligo dT (16 residues) linked to an adaptor primer
b3 (AGAATTCTGCATGCCATGGTCGACGAAGC(T)₁₆). Amplification was performed using the adaptor primer (AGAATTCTGCATGCCATGGTCGACG) and an internal primer (GGCTACGCCATGAACTTCTGCATA) of the MP-121 sequence. The amplification products were reamplified using a nested internal primer (ACATAGCAGGCATGCCTGGTATTG) of the MP-121 sequence and the adaptor primer. The reamplification products were cloned after restriction with Sph I in the likewise restricted vector pT7/T3 U19 (Pharmacia) and sequenced with the sequencing kit "Sequenase Version 2.0" (United States Biochemical Corporation). Clones were characterized by their sequence overlap to the 3' end of the known MP-121 sequence.

One clone, called p121Lt 3' MP13, was used to isolate a NcoI (blunt ended with T4 polymerase)/SphI fragment. This fragment was ligated into a pSK MP-121 (OD/OID) vector, where the OD primer sequence was located close to the T7 primer sequence of the pSK+ multiple cloning site, opened with SphI/SmaI. The resulting plasmid was called pMP121DFus6. It contains MP-121 specific sequence information starting from position 922 and ending with position 1360 of SEQ ID NO. 2.

- 1.9 Using a DdeI fragment of pMP-121DFus6 as a probe, ranging from nucleotide 931 to nucleotide 1304 of SEQ ID NO. 2, a human liver cDNA library (Clontech, # HL3006b, Lot 36223) was screened by a common method described in

detail by Ausubel et al. (Current Protocols in Molecular Biology, published by Greene Publishing Associates and Wiley-Interscience (1989)). From 8.1×10^5 phages, 24 mixed clones were isolated and re-screened using the DdeI fragment. 10 clones were confirmed and the EcoRI fragments subcloned into Bluescript SK (Stratagene, # 212206). EcoRI restriction analysis showed that one clone (SK121 L9.1, deposited by the DSM (#9177) has an insert of about 2.3 kb. This clone contains the complete reading frame of the MP121 gene and further information to the 5' and 3' end in addition to the sequence isolated from mRNA by the described amplification methods. The complete sequence of the EcoRI insert of SK121 L9.1 is shown in SEQ ID NO.2. The reading frame of the MP-121 gene could be confirmed by sequencing of another clone (SK121 L11.1), having the identical reading frame sequence as SK121 L9.1. The beginning of the start codon of the MP-121 sequence of SK121 L9.1 could be determined at position 128 of SEQ ID NO.2, since there are three stop codons in-frame in front of the start codon at positions 62, 77 and 92. The start site of the mature MP-121 is at position 836 of SEQ ID NO.2 in sequence analogy to other members of the TGF- β -family, corresponding to amino acid 237 in SEQ ID NO.4. The stop codon is at position 1184 of SEQ ID NO.2.

Plasmid SK121 L9.1 was deposited under number 9177 at DSM (Deutsche Sammlung von Mikroorganismen und Zellkulturen), Mascheroder Weg 1b, Braunschweig, on 26.04.94).

Example 2
Isolation of MP-52

A further cDNA sequence, MP-52, was isolated according to the above described method (Example 1) by using RNA from human embryo (8-9 weeks old) tissue. The PCR reaction contained cDNA corresponding to 20 ng of poly (A⁺)RNA as starting material. The reamplification step was repeated twice for both enzyme combinations. After ligation, 24 clones from each enzyme combination were further analyzed by sequence analysis. The sample restricted by AlwN I and Sph I yielded a new sequence which was named MP-52. The other clones comprised mainly BMP6 and one BMP7 sequence. The sample restricted by Ava I, AlwN I and Tfi I contained no new sequences, but consisted mainly of BMP7 and a few Inhibin β A sequences.

Goldberg
~~The clone was completed to the 3' end according to the above described method (Example 1). The same embryo mRNA, which was used for the isolation of the first fragment of MP-52, was reverse transcribed as in Example 1. Amplification was performed using the adaptor primer (AGAATTCTGCATGCCATGGTCGACG) and an internal primer (CTTGAGTACGAGGCTTCCACTG) of the MP-52 sequence. The amplification products were reamplified using a nested adaptor primer (ATTCGCATGCCATGGTCGACGAAG) and a nested internal primer (GGAGCCCACGAATCATGCAGTCA) of the MP-52 sequence. The reamplification products were cloned after restriction with Nco I in a likewise restricted vector (pUC 19 (Pharmacia #27-4951-01) with an altered multiple cloning site containing a unique Nco I restriction site) and sequenced. Clones were characterized by their sequence overlap to the 3' end of the known MP-52 sequence. Some of these clones contain the last 143 basepairs of the 3' end of the sequence shown in SEQ ID NO: 1 and the 0,56 kb 3' non translated region (sequence not shown). One of these was used~~

as a probe to screen a human genomic library (Stratagene #946203) by a common method described in detail by Ausubel et al. (Current Protocols in Molecular Biology, published by Greene publishing Associates and Wiley-Interscience (1989)). From 8×10^5 λ phages one phage (λ 2.7.4) which was proved to contain an insert of about 20 kb, was isolated and deposited by the DSM (#7387). This clone contains in addition to the sequence isolated from mRNA by the described amplification methods sequence information further to the 5' end. For sequence analysis a Hind III fragment of about 7,5 kb was subcloned in a likewise restricted vector (Bluescript SK, Stratagene #212206). This plasmid, called SKL 52 (H3) MP12, was also deposited by the DSM (# 7353). Sequence information derived from this clone is shown in SEQ ID NO: 1. At nucleotide No. 1050, the determined cDNA and the respective genomic sequence differ by one basepair (cDNA: G; genomic DNA: A). We assume the genomic sequence to be correct, as it was confirmed also by sequencing of the amplified genomic DNA from embryonic tissue which had been used for the mRNA preparation. The genomic DNA contains an intron of about 2 kb between basepairs 332 and 333 of SEQ ID NO: 1. The sequence of the intron is not shown. The correct exon/exon junction was confirmed by sequencing an amplification product derived from cDNA which comprises this region. This sequencing information was obtained by the help of a slightly modified method described in detail by Frohman (Amplifications, published by Perkin-Elmer Corporation, issue 5 (1990), pp 11-15). The same embryo RNA which was used for the isolation of the 3' end of MP-52 was reverse transcribed using an internal primer of the MP-52 sequence oriented in the 5' direction (ACAGCAGGTGGTGGTGTGGACT). A polyA tail was appended to the 5' end of the first strand cDNA by using terminal transferase. A two step amplification was performed first by application of a primer consisting of oligo dT and an adaptor primer, (AGAATTCTGCATGCCATGGTCGACGAAGC(T_{16})) and secondly an

~~Ay
cont~~ adaptor primer (AGAATTCGCATGCCATGGTCGACG) and an internal primer (CCAGCAGCCCATCCTTCTCC) of the MP-52 sequence. The amplification products were reamplified using the same adaptor primer and a nested internal primer (TCCAGGGCACTAATGTCAAACACG) of the MP-52 sequence. Consecutively the reamplification products were again reamplified using a nested adaptor primer (ATTCGCATGCCATGGTCGACGAAG) and a nested internal primer (ACTAACGTCAAACACGTACCTCTG) of the MP-52 sequence. The final reamplification products were blunt end cloned in a vector (Bluescript SK, Stratagene #212206) restricted with EcoRV. Clones were characterized by their sequence overlap to the DNA of ~~A~~ 2.7.4.

Plasmid SKL 52 (H3) MP12 was deposited under number 7353 at DSM (Deutsche Sammlung von Mikroorganismen und Zellkulturen), Mascheroder Weg 1b, 3300 Braunschweig, on 10.12.1992.

Phage ~~A~~ 2.7.4. was deposited under number 7387 at DSM on 13.1.1993.